MAPPING FOREST COMPOSITION IN THE CENTRAL APPALACHIANS USING AVIRIS: EFFECTS OF TOPOGRAPHY AND PHENOLOGY

Jane R. Foster and Philip A. Townsend[1]

1. INTRODUCTION

Of the many promising applications of imaging spectroscopy in forest ecosystems is the possibility of mapping species composition and distribution with greater accuracy than is possible using standard multispectral data. In areas with complex forest composition, vegetation mapping with hyperspectral imagery requires methodological approaches that can handle the complex spectral and vegetation data sets. In addition, it requires the careful implementation of image preprocessing routines that reduce the effects of other surface properties that may confound the mapping of forest composition. In this paper, we outline an approach to mapping the distribution of three prominent species in the central Appalachian Mountains using AVIRIS imagery and classification/regression trees. In addition, we evaluate the influence of terrain characteristics on mapping results, and finally compare the utility of spring vs. summer AVIRIS imagery for forest mapping.

2. STUDY AREA AND METHODS

2.1 Study Area

The study area is the 15,700 ha Green Ridge State Forest (GRSF) in western Maryland (Figure 1). GRSF is located in the Ridge and Valley physiographic province of the central Appalachian Mountains, and is characterized by steep mountains with deep valleys. Elevation ranges from 200-700 m. The forests were largely cleared around the turn of the twentieth century, and are now mostly intact and mature. Some scattered clearcuts and selective cuts are located within the study area, but are limited in extent. A few burned areas are also present; however, the largest areas of uneven forest are located in places where gypsy moth defoliation over the last two decades has caused substantial tree mortality. Forests are comprised largely of deciduous oaks, with Virginia pine (*Pinus virginiana*) on some west-facing slopes and hemlock (*Tsuga canadensis*) in some valley bottoms. The key species include white oak (*Quercus alba*), especially on mesic slopes and at lower elevations, red oak (*Quercus rubra*), and chestnut oak (*Quercus prinus*) on ridge tops, high elevations and rocky slopes. The understory is largely open, although blueberry (*Vaccinium spp*.) and the evergreen mountain laurel (*Kalmia latifolia*) can be locally abundant.

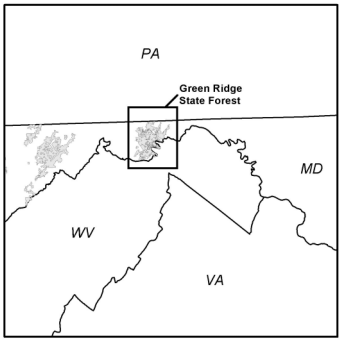

Figure 1. Location of Green Ridge State Forest in western Maryland.

2.2 Field Data

Continuous Forest Inventory (CFI) data from the Maryland Department of Natural Resources was used to characterize forest composition. The CFI database includes 436 plots in GRSF, all of which were sampled in 2000 or 2001. Each CFI plot is a 0.2 acre (0.08 ha) circular area on which all trees > 12 cm diameter are identified and measured. The CFI plots are arrayed on a regular grid at approximately 550 m intervals, yielding a statistical sample of the population of forest properties within the study area. We processed the tree diameter data to determine total basal area (TBA) of each plot, as well as basal area by species (SBA) and relative basal area by species (RBA) for each plot. We use RBA as a measure of species abundance at a plot level. All plots were geographically referenced using a Trimble Pathfinder Pro GPS. For this research, we concentrate on the distribution of three of the most common species in the study area: white oak, chestnut oak, and Virginia pine.

[1] University of Maryland Center for Environmental Science, Appalachian Laboratory, Frostburg, Maryland 21532; jfoster@al.umces.edu

2.3 Image Data and Preprocessing

The research presented here uses two AVIRIS images, acquired on 14 May 2000 and 13 July 2001 from an altitude of ~19,900 m (Table 1). The images were atmospherically corrected using the ATmospheric REMoval program (ATREM) (Gao et al. 1993). The AVIRIS imagery exhibited a cross-track view-angle dependent brightness gradient. This gradient of increasing brightness on the west side of the images results from the AVIRIS scan angle and direction, flight path orientation and solar azimuth, and was corrected by fitting a first-order additive quadratic curve to the mean radiance by view angle (Kennedy et al. 1997). The images were georeferenced to UTM coordinates in ENVI using a triangulation method with > 70 GCPs per scene and nearest neighbor resampling. The images exhibited the effects of differential terrain illumination due to the presence of steep north-south trending mountain ridges in the study area. We applied an empirical correction method (Allen 2000, Meyer et al. 1993), which is described in detail elsewhere in this proceedings (Townsend and Foster 2002). One objective of the research presented here is to determine the consequences of terrain normalization on the capability to map forest composition. Mean spectra for a 3 by 3 pixel window around each plot was collected for all plots in the AVIRIS flight line. This resulted in a sample of 331 plots for the 5/14/2000 scene (Figure 2) and 175 plots for the 7/13/2001 scene.

Table 1. Image Characteristics

Sensor	Date	Time (UTC)	Solar Azimuth	Solar Elevation
AVIRIS	5/14/2000	15:42:46	133.8	62.62
AVIRIS	7/13/2001	15:47:47	134.94	66.72

2.4 Image Classification

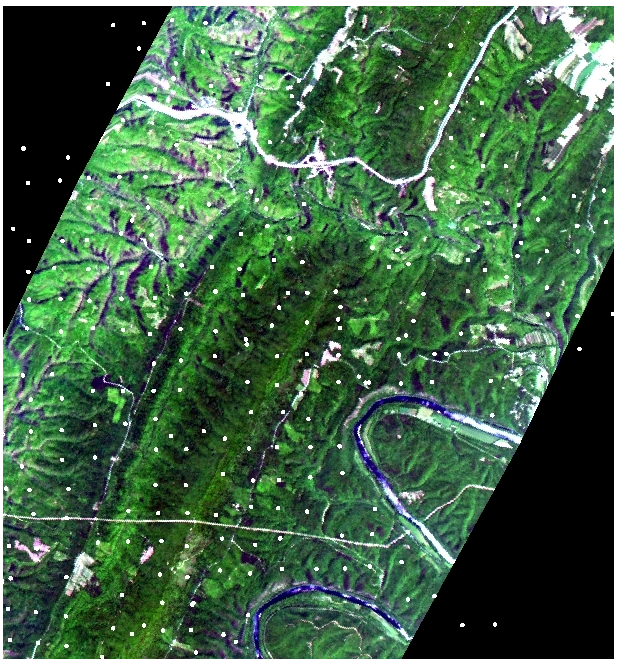

Figure 2. True color 14 May 2000 AVIRIS image of Green Ridge State Forest. White dots indicate locations of CFI plots.

The forests in GRSF exhibit a range of structure and density, with total basal area ranging from recently clear-cut (0) to 60 m^2ha^{-1}. The majority of the forests have BA values between 16-32 m^2ha^{-1}, which is typical for 100-year old forests in the region. However, because gross differences in vegetation structure influence overall reflectance from forested plots, we constrained our analyses to mature, fully stocked forests, i.e., those with TBA > 20 m^2ha^{-1}. This stratification required mapping TBA to delineate forests based on BA; for this, we used multiple stepwise regression (following Townsend 2002) to map TBA as a function of AVIRIS image bands. Following this, we used classification and regression trees (CART) to map composition of individual species. A variety of approaches have been used to map forest composition with hyperspectral data in mountainous landscapes (Martin et al. 1998) and with mixed forests (van Aardt and Wynne 2001), including discriminant analysis and maximum likelihood classifiers. CART is being used increasingly for mapping from remotely sensed imagery (Friedl and Brodley 1997, Friedl et al. 1999, Hess et al. 1995, Simard et al. 2000) and as such is only generally described here. Classification and regression trees (also known as decision trees) are fitted by binary recursive partitioning, in which data sets are consecutively divided into smaller subsets with increasing statistical homogeneity (Clark and Pregibon 1993). Classification trees are used with class data, while regression trees are used to predict continuous data. CART approaches are desirable because they are less sensitive to non-linearities in the input data than methods that require assumptions of Gaussian distributions (as do many image classification techniques) (Clark and Pregibon 1993, Venables and Ripley 1994). In addition, CART is an extremely valuable approach for data exploration when a potentially very large set of independent predictor variable are available, e.g. with hyperspectral data sets. CART can be used to determine the best set of bands for predicting cover characteristics and does not require data reduction, tests for normality or data transformations. One limitation

to CART is that its performance is most robust and repeatable with large data sets, such as the CFI vegetation database that we used. For this work, we used classification trees with relative basal area by species grouped into five categories (A = 0%; B = >0 – 10%; C = 10 – 30%; D = 30 – 50%; E = 50 – 100% RBA). These correspond to abundance levels of none, low, medium, moderate and high. The classification tree models were pruned to avoid overfitting the model to the training data using a cost complexity method designed to minimize misclassified results.

3. PRELIMINARY RESULTS

3.1 Total Plot Basal Area

Total Basal Area (TBA) was best modeled in the corrected 5/14/2000 image using a stepwise linear regression involving 13 image bands with an R^2 of 0.356 ($p < 0.03$). Although the R^2 was not as strong as desired, the relationship was deemed strong enough to subset the AVIRIS image for mapping species composition in mature forests (TBA ≥ 20) using CART. The bands employed in the TBA regression included three in the red portion of the spectrum, one on the red edge, and several in the far end of the near infrared. The regression equation was used to map TBA for the AVIRIS scene, and was then smoothed using a 5x5 median filter to assess the appropriateness of the spatial patterns of TBA created by the model (Figure 3). The map showed relatively low levels of plot BA in recently clear-cut areas throughout the forest, and predicted higher total basal area on mesic slopes grading into valley bottoms. This resulting map exhibits promise for mapping BA, especially because the overall histogram for the distribution of BA on the AVIRIS map matches that from the CFI sample (Figure 4). The prediction of BA from the 13 July image is not shown here. Regression using bands from the July image had an $R^2 < 0.2$. The July image was characterized by fully leafed out forests, which probably obscured some of the variation in TBA on that image.

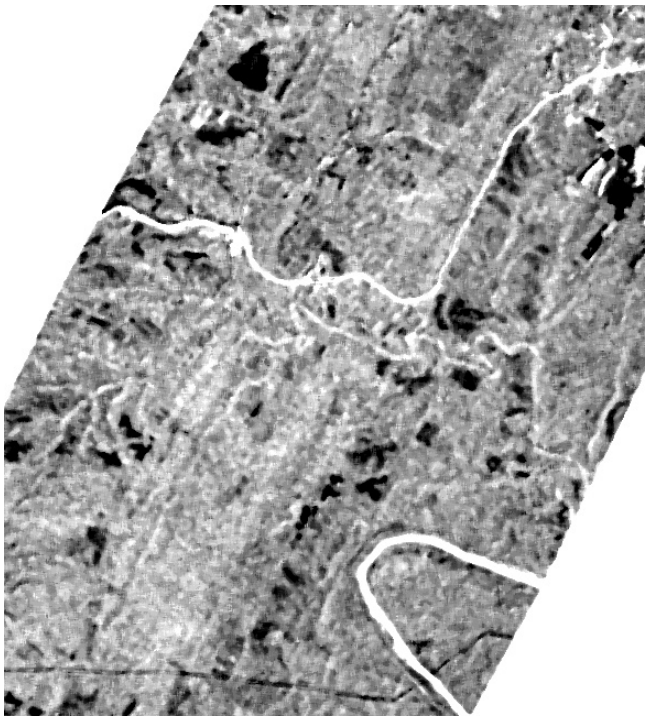

Figure 3. Predicted basal area for forests in Green Ridge State Forest. Gray scale ranges from 0 (black) to > 40 m^2ha^{-1} (white).

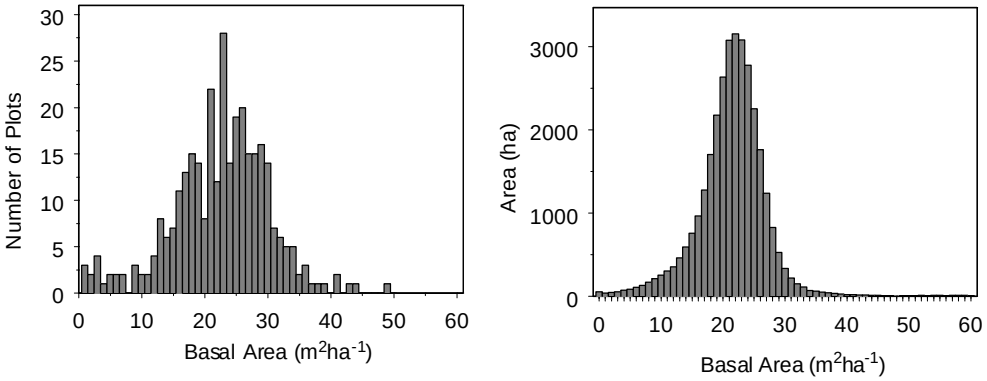

Figure 4. Comparison of the distribution of actual basal area (left) for the CFI plots with predicted BA for the entire study area (right).

3.2 Species Distributions

The results for predicting species abundance classes are listed in Table 2 for each date and for the normalized and non-normalized imagery. The best model for each species is highlighted. An example classification tree for Virginia pine is presented in Figure 5, with the resultant map in Figure 6. The bands used in all of the classification tree models are listed in Table 3.

Table 2. Classification tree results.

	Quercus alba		*Quercus prinus*		*Pinus virginiana*	
Image Date	Normalized	Original	Normalized	Original	Normalized	Original
14 May 2000	0.6009	0.6479	0.6291	0.6479	0.6901	0.6948
13 July 2001	0.6387	**0.6807**	**0.6891**	0.605	0.7479	**0.7563**

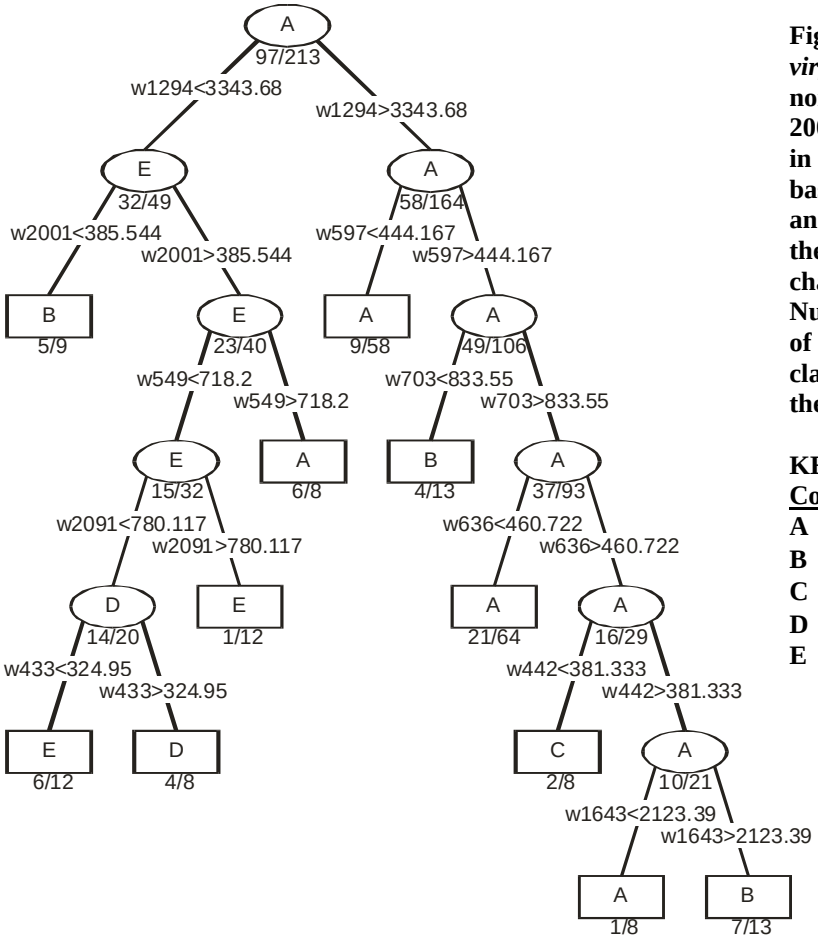

Figure 5. Classification tree for *Pinus virginiana*, generated from the normalized AVIRIS image of 14 May 2000. Each branch delineates a split in the dependent variable (relative basal area of *P. virginiana*) based on an AVIRIS image band (identified by the prefix w, e.g. w1294 = the AVIRIS channel centered on 1294 nm). Numbers below each box (node or leaf of the tree) indicate the mis-classification rate for that branch of the decision tree.

KEY

Code	Description
A	**RBA = 0%**
B	**0% < RBA ≤ 10%**
C	**10% < RBA ≤ 30%**
D	**30% < RBA ≤ 50%**
E	**RBA > 50%**

3.2.1 Effect of Image Date

The classification accuracies for species abundance (i.e., relative basal area) vary from 60-76%. In general, the summer image performed better for predicting the individual species. This probably results from the likelihood that not all of the deciduous trees in the region (including white oak and chestnut oak) were fully leafed out on 14 May 2000, making their detection somewhat problematic. The fact that classification accuracy for Virginia pine was relatively high for both the spring and summer images points to its ease of detection as an evergreen conifer. Note that even though the misclassification rates are high, our classification categories are very specific (*relative*

abundances by species). We consider these results to be very promising for mapping individual species patterns using decision trees. The logical next step to this research is to employ a multitemporal classification to take advantage of the differing reflectance characteristics of each species across seasons.

3.2.2 Topographic Effects

The comparison of models that used terrain normalized imagery versus the uncorrected original imagery yielded mixed results. For chestnut oak, which is typically located along steep slopes and ridge tops, the terrain-normalized imagery produced the best classification of species distribution. This is no surprise, as chestnut oak is typically located on the east and west slopes of the ridges in GRSF. The normalization would therefore be expected to improve the classification of this species by reducing differences in reflectance between illuminated and shadowed slopes. For Virginia pine, there was very little difference between the accuracy of models using normalized imagery and uncorrected imagery. This is likely due to the very distinct reflectance characteristics of Virginia pine, making it easily identifiable regardless of illumination effects. Finally, white oak was best predicted using the uncorrected imagery. This was somewhat of a surprise. However, because white oak generally occurs on gentle slopes, lower slopes and in flat areas, it is possible that the terrain normalization had very little impact on areas where white oak is found, and perhaps introduced some confusion in areas where white oak mixes with other deciduous oaks. Although these results do not provide

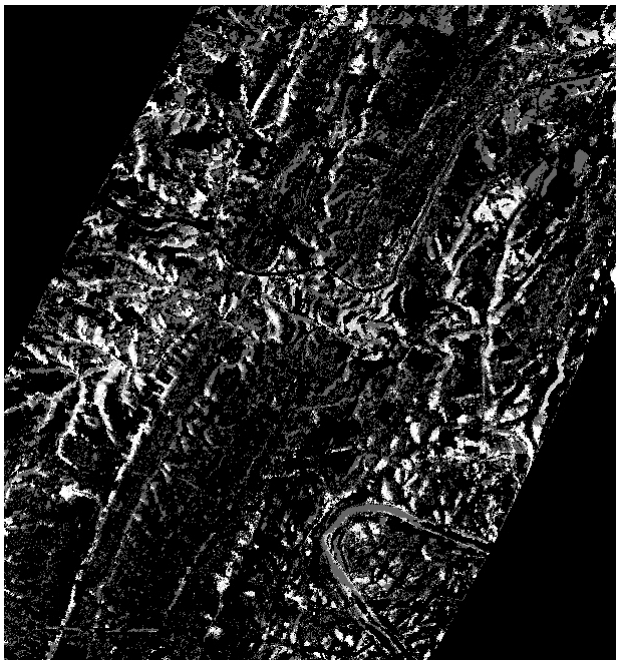

Figure 6. Prediction of relative basal area for *P. virginiana*, ranging from 0% (black areas) to > 50% (white). Virginia pine is especially prominent along west-facing slopes of ridges.

Table 3. Bands used in the classification tree models.

5/14/2000 Normalized			5/14/2000 Original			7/13/2001 Normalized			7/13/2001 Original		
QUAL	**QUPR**	**PIVU**	**QUAL**	**QUPR**	**PIVU**	**QUAL**	**QUPR**	**PIVU**	**QUAL**	**QUPR**	**PIVU**
452	491	433	433	452	423	451	422	471	432	442	422
529	510	442	510	578	500	461	432	490	451	538	693
655	520	539	549	607	655	741	471	538	529	2010	769
665	529	597	703	694	713	807	558	558	731	2060	1117
703	549	636	713	1155	799	1079	664	567	798	2379	1154
1193	568	703	741	1165	1474	1293	702	674	1135		1254
1444	587	1294	1174	1334	1494	1533	865	1333	1572		2010
1484	703	1643	1454	2001	1514	2020	2180	2020	1632		2060
1623	1334	2001	1533	2011	2001	2389	2270	2349	1772		2379
1792	1444	2091	1633	2021	2330	2438		2419	1782		2419
2041	1494		1703	2031					2309		2438
2211	1524		2011	2151					2438		
2420	2011		2031	2380							
	2191		2231								
	2410										
	2430										

QUAL = *Quercus alba* (white oak)
QUPR = *Quercus prinus* (chestnut oak)
PIVU = *Pinus virginiana* (Virginia pine)

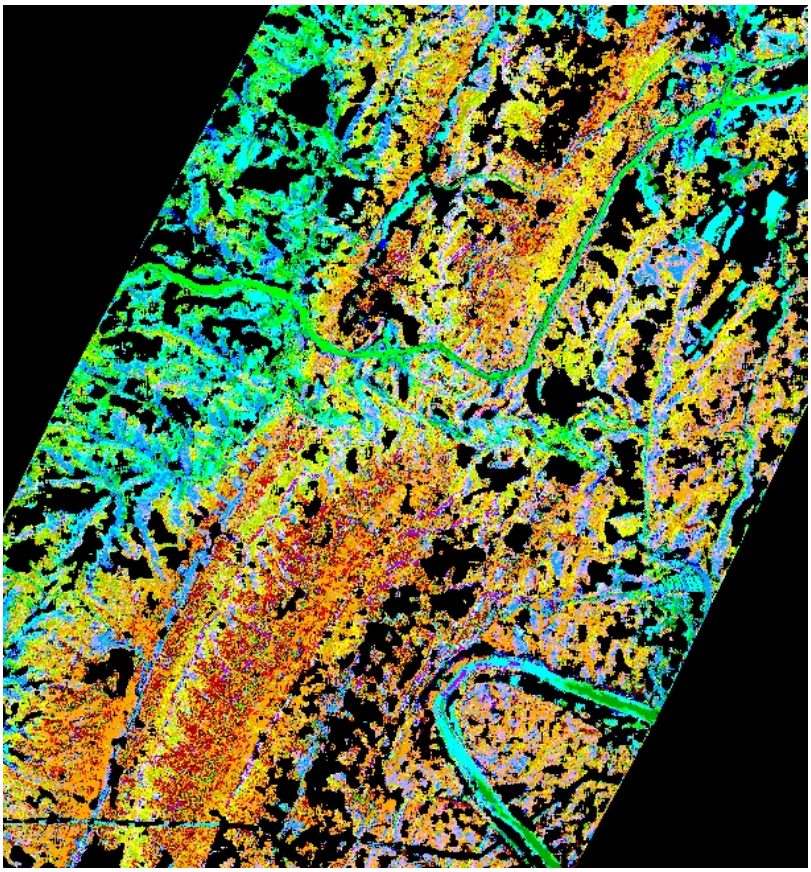

Figure 7. Composite prediction of the distributions of Virginia pine (projected in the blue channel), white oak (green channel), and chestnut oak (red channel). Black areas have TBA < 20 m^2ha^{-1}; the river and interstate are not masked out on this image. Red areas represent ridges and steep slopes dominated by chestnut oak; blue areas are west-facing slopes dominated by pines (and other conifers); green areas are mesic slopes dominated by white oak. Other colors represent mixes.

unequivocal support for the need to implement topographic normalizations, our results do suggest that more research is needed to determine the relative value of terrain normalization for species mapping.

4. CONCLUSIONS

A composite map of species distributions generated from the classification trees yields a map (Figure 7) that closely resembles the actual distributions and dominance of the three species described in this paper. These results represent a promising first step towards the application of AVIRIS imagery for detailed species mapping in rugged terrain. The results were mixed regarding the value of the terrain normalization; however, we expect that a multi-temporal analysis of the normalized AVIRIS imagery will result in improved classification results for all species. Finally, decision trees represent a promising technique for mapping using complex data sets (such as the CFI field data and AVIRIS hyperspectral data), reducing the need for data reduction or data transformation.

5. ACKNOWLEDGMENTS

This research was supported by NASA EO-1 Science Validation Team grant NCC5-493. Thanks to Robert Chastain and Brian Sturtevant for their contributions to this work, and to Chris McCann and the Maryland DNR CFI program for contributing forest data. This paper is Scientific Series Contribution Number 3551-AL, University of Maryland Center for Environmental Science.

6. REFERENCES CITED

Allen, T. R. 2000. Topographic normalization of Landsat Thematic Mapper data in three mountain environments. *Geocarto International* 15: 13-19.

Clark, L. A., and D. Pregibon. 1993. Tree-based models. Pages 377-419 in J. M. Chambers and T. J. Hastie, eds., *Statistical Models in S.* Chapman and Hall, New York.

Friedl, M. A., and C. E. Brodley. 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment* 61: 399-409.

Friedl, M. A., C. E. Brodley, and A. H. Strahler. 1999. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing* 37: 969-977.

Gao, B. C., K. B. Heidebrecht, and A. F. H. Goetz. 1993. Derivation of scaled surface reflectances from AVIRIS data. *Remote Sensing of Environment* 44: 165-178.

Hess, L. L., J. M. Melack, S. Filoso, and Y. Wang. 1995. Delineation of inundated area and vegetation along the Amazon floodplain with the SIR-C synthetic aperture radar. *IEEE Transactions on Geoscience and Remote Sensing* 33: 896-904.

Kennedy, R. E., W. B. Cohen, and G. Takao. 1997. Empirical methods to compensate for a view-angle-dependent brightness gradient in AVIRIS imagery. *Remote Sensing of Environment* 62: 277-291.

Martin, M. E., S. D. Newman, J. D. Aber, and R. G. Congalton. 1998. Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment* 65: 249-254.

Meyer, P., K. I. Itten, T. Kellenberger, S. Sandmeier, and R. Sandmeier. 1993. Radiometric corrections of topographically induced effects on Landsat TM data in an Alpine environment. *ISPRS Journal of Photogrammetry and Remote Sensing* 48: 17-28.

Simard, M., S. S. Saatchi, and G. De Grandi. 2000. The use of decision tree and multiscale texture for classification of JERS-1 SAR data over tropical forest. *IEEE Transactions on Geoscience and Remote Sensing* 38: 2310-2321.

Townsend, P. A. 2002. Estimating forest structure in wetlands using multi-temporal SAR. *Remote Sensing of Environment* 79: 288-304.

Townsend, P. A., and J. R. Foster. 2002. Terrain normalization of AVIRIS and Hyperion imagery in forested landscapes. In R. O. Green, ed., *Proceedings of the Eleventh JPL Airborne Earth Science Workshop*. JPL, Pasadena, California.

van Aardt, J. A. N., and R. H. Wynne. 2001. Spectral separability among six southern tree species. *Photogrammetric Engineering and Remote Sensing* 67: 1367-1375.

Venables, W. N., and B. D. Ripley. 1994. *Modern Applied Statistics with S-Plus.* Springer-Verlag, New York.